Acta Crystallographica Section F
**Structural Biology and Crystallization Communications**
ISSN 1744-3091

# Structure of pig heart citrate synthase at 1.78 Å resolution

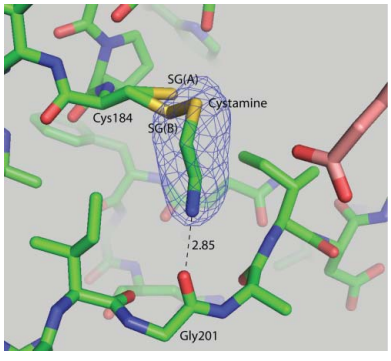
**Steven B. Larson,[a] John S. Day,[a] Chieugiang Nguyen,[b] Robert Cudney[b] and Alexander McPherson[a]***

[a]Department of Molecular Biology and Biochemistry, The University of California, Irvine, CA 92697-3900, USA, and [b]Hampton Research, Aliso Viejo, CA 92656-3317, USA

Correspondence e-mail: amcphers@uci.edu

Received 30 January 2009
Accepted 7 March 2009

**PDB Reference:** pig heart citrate synthase, 3enj, r3enjsf.






Pig heart citrate synthase was crystallized from a small-molecule cocktail containing cystamine dihydrochloride, aspartame and benzamidine hydrochloride. The structure was refined to an $R$ factor of 0.179 ($R_{\rm free}$ = 0.222) using synchrotron data to a resolution of 1.78 Å. The model includes the full-length protein, a chloride ion, a sulfate ion, 305 water molecules and an unexpected moiety attached through a disulfide linkage to Cys184, which was modeled as a half-cystamine molecule generated by disulfide exchange with the cystamine in the small-molecule cocktail.

## 1. Introduction

The enzyme citrate synthase catalyzes what is considered to be the first step in the citric acid or Krebs cycle proper. Its function is to transfer the acetate group of acetyl-CoA to oxaloacetate, producing citrate. The enzyme from pig heart is a dimeric protein composed of two identical subunits. Each monomer is comprised of 437 amino acids and has a molecular weight of about 50 kDa. The monomer is divided into a large and small domain separated by a deep cleft that forms the active site of the enzyme. Ordered binding of oxaloacetate and acetyl-CoA is accompanied by rotation of one protein domain to effectively close the cleft and preclude the entry of water.

The structure of the pig heart enzyme was reported by Remington *et al.* (1982) (PDB codes 1cts and 2cts) and Wiegand *et al.* (1984) (PDB code 4cts). Although no additional models of this specific enzyme have been deposited in the PDB subsequently, a number of structures of citrate synthase from other organisms in both the open and the closed conformation and in ternary complexes with substrate analogues have been determined.

In the course of our investigations (McPherson & Cudney, 2006; Larson *et al.*, 2007, 2008) on the propensity of mixtures, or 'cocktails', of conventional small molecules to promote the crystallization of macromolecules, we used pig heart citrate synthase as one of our test proteins. It was particularly useful because although it did crystallize, we found it difficult to crystallize under the conditions that we were using. It frequently formed microcrystals or crystals that were otherwise unsuitable for X-ray diffraction and only rarely (as we describe below) yielded crystals that diffracted well. In these studies, pig heart citrate synthase provided a notable success for the experimental trials. This X-ray diffraction analysis of the crystals suggests the likely source of this result.

## 2. Materials and methods

Pig heart citrate synthase was purchased from Sigma Biochemical Co. (St Louis, Missouri, USA) as an ammonium sulfate suspension and was dialyzed against water to a concentration of about 8 mg ml$^{-1}$. The dialysis procedure resulted in the formation of a substantial amount of precipitate, which was removed by centrifugation. The screening experiments that yielded the citrate synthase crystals investigated in this work have been described in detail in another paper (Larson *et al.*, 2008). The crystals were grown by vapor diffusion at room temperature in 96-well Intelli-Plates (Hampton Research, Aliso Viejo, California, USA) with 90 µl reservoirs of 25%(*w*/*v*) PEG 3350 in water. The protein droplets were of 4 µl

**Table 1**
Data-processing, refinement and model statistics.

Values in parentheses are for the highest resolution shell.

| | |
|---|---|
| Crystal data | |
| Space group | $P4_12_12$ |
| $Z$ | 8 |
| Unit-cell parameters (Å) | $a = b = 76.34$, $c = 198.15$ |
| Data-processing statistics | |
| Resolution (Å) | 41.56–1.72 (1.78–1.72) |
| No. of unique reflections | 63039 (6200) |
| Redundancy | 6.35 (6.07) |
| $\langle I/\sigma(I)\rangle$ | 7.3 (1.4) |
| Completeness | 99.8 (100.0) |
| $R_{\rm merge}$† | 0.072 (0.646) |
| Structure-refinement statistics | |
| Resolution (Å) | 41.56–1.78 |
| No. of reflections ($F_o = 0$ removed) | 49141 |
| $R/R_{\rm free}$‡ (all data) | 0.179/0.222 |
| Reflections in test set (%) | 5.10 |
| No. of refined parameters | 15434 |
| No. of reflections (working set) | 46514 |
| No. of restraints | 56077 |
| Data-to-parameter ratio | 3.01 |
| Data/restraints-to-parameter ratio | 6.65 |
| Model statistics | |
| Non-H atoms | |
| Protein atoms (full/partial) | 3361/176 |
| Ligand atoms (full/partial) | 1/9 |
| Water atoms (full/partial) | 277/28 |
| Geometry: r.m.s. deviations from ideal values | |
| Bonds (Å) | 0.011 |
| Angles (°) | 1.495 |
| Planes (Å) | 0.009 |
| Chiral centers (Å$^3$) | 0.091 |
| Average isotropic $B$ factors (Å$^2$) | |
| Overall | 50.9 |
| Protein (full/partial) | 50.0/57.3 |
| Ligands (full/partial) | 45.6/48.8 |
| Water (full/partial) | 53.5/38.9 |
| Ramachandran plot | |
| Most favored region | 357 [93.9%] |
| Allowed region | 21 [5.5%] |
| Generously allowed region | 2 [0.5%] |
| Disallowed regions | 0 [0.0%] |

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the $i$th used observation for unique $hkl$ and $\langle I(hkl)\rangle$ is the mean intensity for unique $hkl$. ‡ $R = \sum_{hkl}|F_o - F_c|/\sum_{hkl} F_o$, where $F_o$ and $F_c$ are the observed and calculated structure-factor amplitudes.

volume and consisted of two parts of the stock protein solution and one part each of the 'cocktail' of small molecules and the reservoir solution. All solutions were buffered to pH 7.8 with 0.1 *M* Tris–HCl. The 'cocktail' was made to concentrations of approximately 1%(*w*/*v*) of each of the small molecules, namely aspartame, benzamidine hydrochloride and cystamine dihydrochloride (see Fig. 1). The concentrations of the small molecules in the crystallization drop were approximately 5–10 m*M*. Crystals usually appeared and grew to full size within one to two weeks.

The crystals were captured in cryoloops (Hampton Research, Aliso Viejo, California, USA), flash-frozen in liquid nitrogen and shipped by Federal Express to SSRL. X-ray diffraction data were collected to 1.72 Å resolution by remote procedures on beamline 7-1. The data were integrated, processed, scaled and merged using the program *d*TREK* (Pflugrath, 1999). The space-group symmetry, unit-cell parameters and relevant data-collection statistics are presented in Table 1.

The initial model was obtained by direct model placement in the unit cell of PDB entry 1cts (Remington *et al.*, 1982; Berman *et al.*, 2000) with the citrate ion removed; this is the only citrate synthase structure that is isomorphous with our cell. Refinement was begun with rigid-body refinement followed by simulated annealing using the program *CNS* (Brünger *et al.*, 1998). Iterative cycles of refinement and model rebuilding, including the addition of water, ions and an appendage to Cys184 discussed below, were performed. The program *REFMAC*5 (Murshudov *et al.*, 1997) was ultimately used for refinement in order to take advantage of its TLS capabilities. The structure was divided into two TLS groups corresponding to the small (residues 1–54 and 411–437) and large (residues 55–410) domains of the protein. Otherwise, all non-H atoms were refined isotropically despite the appearance of ANISOU records in the deposited coordinate file 3enj. H atoms in riding positions were included in the model for better treatment of van der Waals contacts during refinement. Partial occupancies of alternate conformations of amino acids were determined by a series of refinements in which the occupancy ratio of the conformations of the disordered residue was varied; the final ratio was selected based upon the lowest $R_{\rm free}$ values and the similarity of $B$ factors for equivalent atoms of the disordered conformations. The final refinement statistics are shown in Table 1.

Model building was performed with the programs *O* (Jones & Kjeldgaard, 1994) and *Coot* (Emsley & Cowtan, 2004). $F_o - F_c$ difference Fourier maps were computed with *CNS* or from MTZ files from *REFMAC5* using *FFT* (Ten Eyck, 1973) in the *CCP*4 program suite (Collaborative Computational Project, Number 4, 1994). Models were displayed and figures were produced with *PyMOL* (DeLano, 2002). The RCSB validation server (Berman *et al.*, 2000) and *PROCHECK* (Laskowski *et al.*, 1993) were used to assess the quality of the model.

## 3. Results

The final citrate synthase model contained 437 amino acids (the full sequence), a chloride ion in the active site, a sulfate ion, 305 full or partial water molecules and a half cystamine (at 75% occupancy) attached to the thiol of Cys184. Including Cys184, there are 15 discretely disordered residues in the protein model, 17 partially occupied water molecules and a partially occupied sulfate ion. Although all residues were included in the model and were refined, the electron density at the N-terminus (especially Ala1, whose occupancy was set to 0.50) was not very well defined. Likewise, the electron density at the C-terminus was weak. The loop formed by

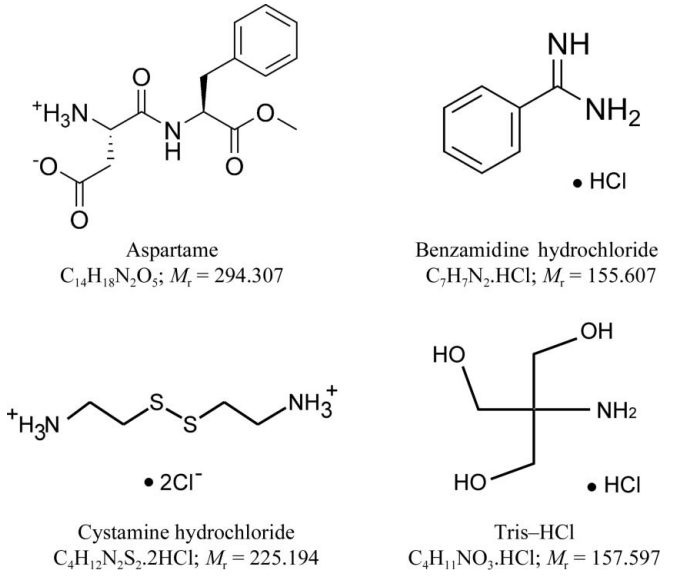



**Figure 1**
Schematic diagrams of the structures of the small molecules in the 'cocktail' and buffer used in this study.

residues 362–368 has very high $B$ values. Residues Glu239 and Lys366 fall in the generously allowed region of the Ramachandran plot. The final model gave $R$ = 0.179 and $R_{free}$ = 0.222 for data to 1.78 Å resolution. Refinement and model statistics are given in Table 1.

Inspection of the $F_o - F_c$ electron-density difference map revealed density that was contiguous with the thiol group of Cys184 (shown in Fig. 2). The elongated density extended away from the thiol towards the carbonyl O atom of Gly201 of the same molecule. Our interpretation, consistent with the chemistry of cystamine and the process of disulfide exchange, is that half of a cystamine molecule formed a disulfide bond with Cys184 on the surface of the protein. By adding the half-cystamine molecule (S—C—C—N) to the citrate synthase model, we were able to fit the extended density quite well, with the amino group forming a hydrogen bond to the carbonyl O atom of Gly201 (N···O distance of 2.85 Å). However, a fully occupied half-cystamine moiety produced a negative hole at the S atom, which disappeared when the occupancy for this moiety was reduced to 75%.

Besides the expected water structure built up from the difference density Fourier maps, there were three features that were inconsistent with water structure. In the substrate-binding site, adjacent to Arg401, which binds citrate in 1cts and 2cts and oxaloacetate in 4cts, we modeled density as a chloride ion after determining that a water

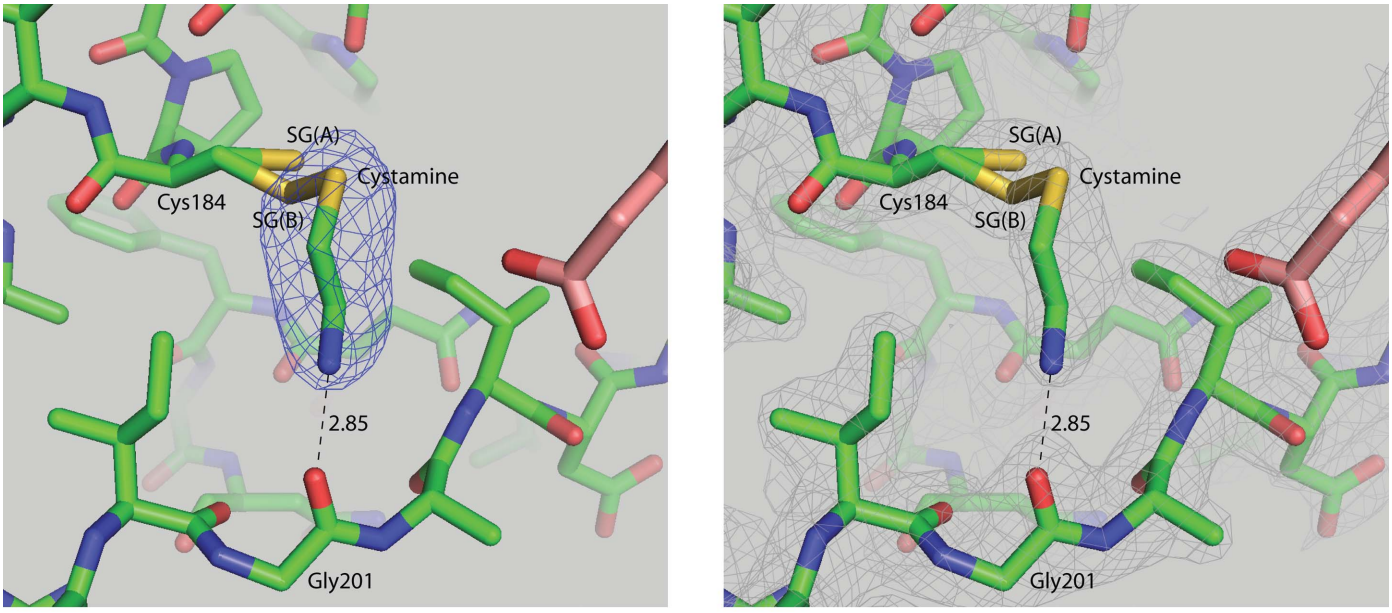



**Figure 2**
(*a*) $F_o - F_c$ difference electron density at 4.0$\sigma$ extending from the Cys184 thiol group with model superimposed. (*b*) Electron density (contoured at 1.0$\sigma$) of a $2F_o - F_c$ map with model superimposed. The N atom of the half-cystamine moiety forms a hydrogen bond to the carbonyl O atom of Gly201.

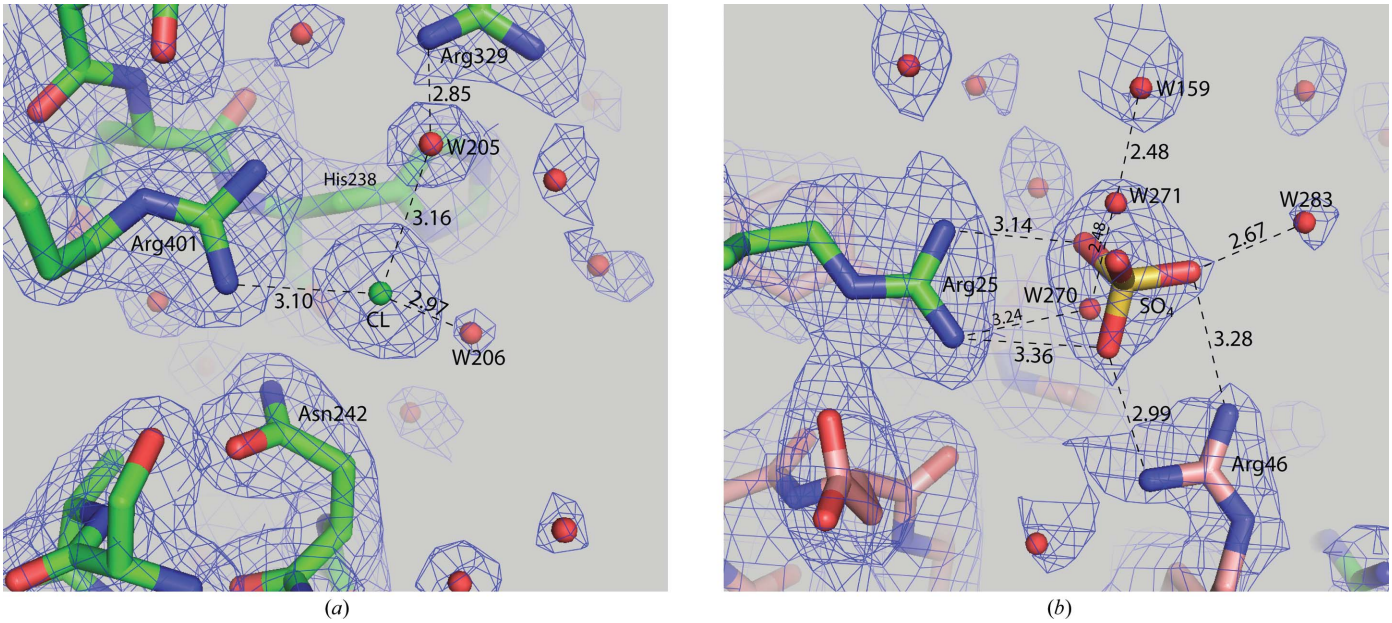



**Figure 3**
(*a*) The $Cl^-$ ion bound to Arg401 with $2F_o - F_c$ electron density overlaid at a contour level of 1.2$\sigma$. (*b*) The $SO_4^{2-}$-ion site bridging between Arg25 of one molecule and Arg46 of a dimer-related molecule. The $2F_o - F_c$ electron density is superimposed at a contour level of 0.8$\sigma$. This site is shared with water at a 50:50 ratio.

**Table 2**
Cocktails used with citrate synthase.

| Cocktail No. | Components |
|---|---|
| 1 | Anthrone, phenylglyoxal, cystamine dihydrochloride |
| 76 | Cystathionine, L-citrulline, cystamine dihydrochloride |
| 151 | Cystamine dihydrochloride, aspartame, benzamidine hydrochloride |
| 181 | Nicotinamide, cystamine dihydrochloride, 1,2-diaminocyclohexane sulfate |

molecule did not adequately account for the density. There were three sources of chloride ions in the crystallization mixture: two from the small-molecule cocktail and the other from the Tris–HCl buffer. The chloride interactions are illustrated in Fig. 3(*a*).

On the surface of the protein dimer, between the guanadinium groups of Arg25 of one molecule and Arg46 of the other, we found density which we modeled as a composite of a sulfate ion and two water molecules in a 50:50 ratio. Difference maps calculated after inclusion of a sulfate ion suggested that this site was shared by water molecules and was subsequently modeled as such. Despite the dialysis of the protein sample, it appears that sulfate remains in the dimer from the initial ammonium sulfate suspension. The interactions involving the sulfate ion are illustrated in Fig. 3(*b*).

The last feature of the map remains unexplained. This odd-shaped density, shown in Fig. 4 contoured at 3.2$\sigma$, was on the surface near Arg195. However, its shape with respect to the small molecules present and its distance from the protein (>3.5 Å) gave no clues for modeling.

Inspection of difference density Fourier maps revealed a number of discretely disordered amino-acid side chains. These were Gln100, Glu105, Arg195, Gln287, Asn307, Val315, Thr331, His343, Gln383, Arg421, Met425, Lys432, Ser436 and Lys437. All of these are on the surface of the protein except Val315. Arg421 was modeled in three conformations. In addition, Ala1 was assigned an occupancy of 50%; the density was poorly defined and the N-terminus appeared to have several directions. This residue was not included in any previous structure of citrate synthase from this species. Also, as previously mentioned, Cys184 is disordered by virtue of its partial conversion to the half-cystamine derivative. No previously reported structure contained discretely disordered residues.

## 4. Discussion

As reported in Larson *et al.* (2008) for experiment V of their series, citrate synthase yielded fine needle-shaped crystals and clusters of needles for quite a number of trials and using a variety of small-molecule 'cocktails'. For four mixtures of conventional small molecules, however, the small needle-shaped crystals were absent and in their place were large perfect bipyramids and thick plates. The unit cell was virtually the same as for the crystals previously described by Remington *et al.* (1982) (PDB code 1cts). The space group was $P4_12_12$, with unit-cell parameters $a = b = 76.34$, $c = 198.15$ Å. Subsequent experiments in which citrate synthase was combined individually with each of the four 'cocktails' demonstrated that the crystals were readily reproduced. The compositions of the four cocktails are shown in Table 2. Inspection of the compositions of the four mixtures immediately shows cystamine to be a common component and probably the causative agent for the results. It should be noted that PEG was not employed in the crystallization procedures of the previously reported structures, only citrate, sodium/potassium phosphate or Tris–HCl buffers and the substrates, co-enzymes or inhibitors under study, although in the crystallization of 4cts polyvinylpyrrolidone was also added to the crystallization mix.

As noted above, we have extended the resolution of the structure of pig heart citrate synthase to 1.78 Å resolution. This alone is noteworthy, we feel, because the current PDB entry (1cts) for the same unit cell was based on 2.7 Å resolution data. Part of the increase in resolution is undoubtedly the consequence of advances in X-ray sources (synchrotron *versus* rotating-anode generators), but not likely all of it. Hence, we feel that the 'cocktails' were responsible not only for producing crystals amenable for X-ray diffraction analysis, but also for a significant advance in the resolution of the crystals. Furthermore, we have produced a model that includes all residues, whereas the deposited structures for this enzyme in three different space groups are lacking from one to three N-terminal amino acids and the last three residues of the C-terminus.

Generally, our structure is very consistent with 1cts, which is in the open conformation. The structures of 2cts and 4cts are in the closed conformation. Superposition of selected $C^{\alpha}$ atoms (segments 5–56, 65–196, 199–275 and 380–433) resulted in r.m.s. deviations of 0.43, 0.57, 0.59 and 0.58 Å for 1cts, 2cts and the two molecules in 4cts, respectively. Based on this superposition, the major backbone deviation of 3.02 Å between 1cts and our model is in the region 276–379. For structures 2cts and 4cts, the maximum deviations in this same region are 11.12, 10.83 and 10.83 Å, which identifies the region of greatest movement when the enzyme changes from the open to the closed conformation. The helix–loop–helix encompassing the half-cystamine modification (residues 165–218) has an r.m.s. deviations for $C^{\alpha}$-atom fit of 0.33, 0.57, 0.60 and 0.55 Å for 1cts, 2cts and the two molecules in 4cts, respectively. Thus, it appears that the attachment of the half-cystamine molecule at Cys184 does not significantly alter the molecular structure around Cys184.

The results with pig heart citrate synthase provide another example that mixtures of conventional small molecules, in this case mixtures containing the common component cystamine, can enhance

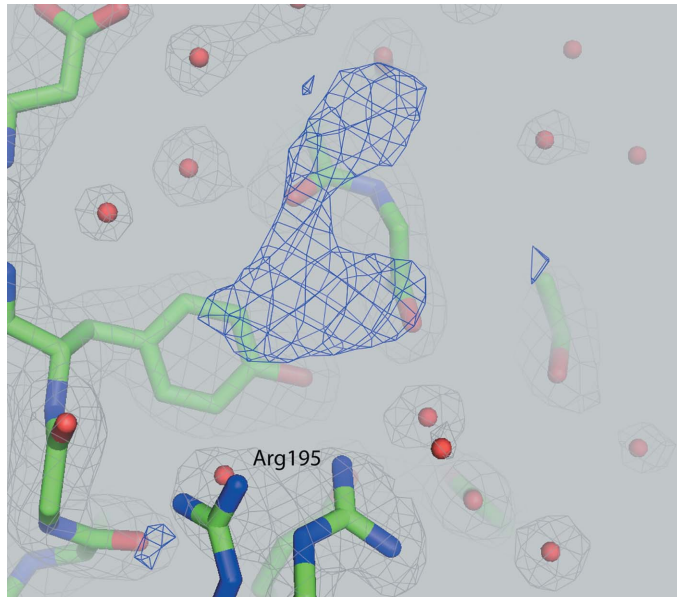



**Figure 4**
The unidentified difference density (blue) adjacent to Arg195 is contoured at 3.2$\sigma$. The shape and size of this density does not conform to any small molecule of the cocktail or buffer and it would appear that any molecule placed in this density would be more than 3.5 Å from Arg195, the closest residue to this density. The $2F_o - F_c$ density, shown in gray, is contoured at 1.1$\sigma$.

the crystallization of macromolecules. It is also significant that the enhancement occurred no matter what the small-molecule partners were in the cocktails. That is, the positive effects of the cystamine were neither dependent upon nor suffered from the presence of the other small molecules. Although we would like to pinpoint the exact reason for the enhanced crystallization and diffraction properties of this cocktail, the facts that the half cystamine on Cys184 neither alters the molecular structure nor engages in intermolecular interactions make it difficult to assign a definitive cause for these enhancements.

This work was supported by NIH Grant GM074899 for the establishment of the Center for High Throughput Structural Biology.

## References

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
DeLano, W. L. (2002). *The PyMOL Molecular Visualization System*. DeLano Scientific, San Carlos, California, USA. http://www.pymol.org.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Jones, T. A. & Kjeldgaard, M. (1994). *O – The Manual*, v.5.10. Uppsala University Press.
Larson, S. B., Day, J. S., Cudney, R. & McPherson, A. (2007). *Acta Cryst.* D**63**, 310–318.
Larson, S. B., Day, J. S., Nguyen, C., Cudney, R. & McPherson, A. (2008). *Cryst. Growth Des.* **8**, 3038–3052.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
McPherson, A. & Cudney, B. (2006). *J. Struct. Biol.* **156**, 387–406.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Pflugrath, J. W. (1999). *Acta Cryst.* D**55**, 1718–1725.
Remington, S., Wiegand, G. & Huber, R. (1982). *J. Mol. Biol.* **158**, 111–152.
Ten Eyck, L. F. (1973). *Acta Cryst.* A**29**, 183–191.
Wiegand, G., Remington, S., Diesenhofer, J. & Huber, R. (1984). *J. Mol. Biol.* **174**, 205–219.